

Lip Synching in Musikclips

Zur Konstruktion von Audio-Vision durch musikbezogene Darstellungshandlungen

Bearbeitet von
Daniel Klug

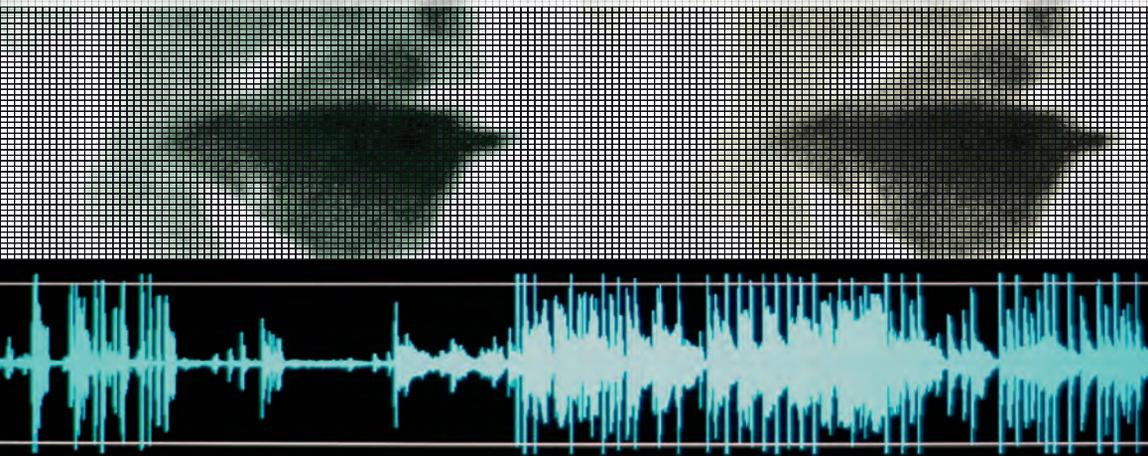
1. Auflage 2013. Buch. 375 S. Kartoniert
ISBN 978 3 8487 0240 4
Gewicht: 549 g

[Recht > Handelsrecht, Wirtschaftsrecht > Urheberrecht, Medienrecht > Medienrecht,
Presserecht, Rundfunkrecht](#)

schnell und portofrei erhältlich bei

**beck-shop.de**
DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.



Daniel Klug

Lip Syncing in Musikclips

Zur Konstruktion von Audio-Vision durch musikbezogene
Darstellungshandlungen



Nomos

Die Reihe „Short Cuts | Cross Media“ nimmt ‚mediale Abkürzungen‘ in den Blick: kurze, daher pointierte und häufig multimodal arbeitende Medienformate und -genres wie etwa den Videoclip, den Werbespot oder Angebote der sog. Mobile Media. Insbesondere auch die von inzwischen radikaler Crossmedialität bzw. Multimodalität geprägten ‚Orte‘ resp. Medien ihrer Präsentation stehen im Mittelpunkt des Interesses – allen voran der radikalste crossmediale/multimodale Präsentationsort, die Website bzw. – kontextuell erweitert – Websites (also die unter einer URL zusammengefassten, multimodalen Angebote des Internet). Geklärt werden soll, ob und inwiefern die Parameter der Kürze und Pointiertheit mithin im Rahmen eines umfassenden Kommunikationsumfelds (crossmediality) zur Beschreibung bestimmter medialer Produkte bzw. Sinneinheiten einen Beitrag zu leisten vermögen und ob sich mit dieser Perspektivierung mitunter ein modernes Produktions- wie Rezeptionsparadigma erkennen und beschreiben lässt, das – gewissermaßen – ‚quer‘ zu anderen, möglichen Beschreibungsdimensionen liegt.

Neben wissenschaftlichen – theoretisch wie method(olog)isch ausgerichteten – Zugängen ist es erklärtes Ziel der Reihe, auch praxisorientierte Arbeiten (etwa aus dem Bereich der Kunst oder Informatik) im Sinne eines Theorie-Praxis-Transfers aufzunehmen.

Schriftenreihe „Short Cuts | Cross Media“

herausgegeben von
Klaus Neumann-Braun,
Axel Schmidt und Henry Keazor

Band 6

Daniel Klug

Lip Synching in Musikclips

Zur Konstruktion von Audio-Vision durch
musikbezogene Darstellungshandlungen



Nomos

@ Titelbild: istockphoto.com

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Basel, Univ., Diss., 2012

ISBN 978-3-8487-0240-4

1. Auflage 2013

© Nomos Verlagsgesellschaft, Baden-Baden 2013. Printed in Germany. Alle Rechte, auch die des Nachdrucks von Auszügen, der fotomechanischen Wiedergabe und der Übersetzung, vorbehalten. Gedruckt auf alterungsbeständigem Papier.

Inhalt

Vorwort	5
Einleitung	11
1 Lip Synching	23
1.1 Lip Synching als audiovisuelles Phänomen	23
1.2 Formatspezifische Definition des Lip Synching im Musikclip	36
1.3 Abgrenzung des Lip Synching im Musikclip zu verwandten audiovisuellen Darstellungen	43
1.3.1 Bauchreden: Ventriloquismus	43
1.3.2 Film: „Moving lips“	46
1.3.3 Ventriloquismus in der (Live-)Performance populärer Musik	49
1.4 Körper, Stimme, Relationen	51
1.5 Lip Synching und Körper	61
1.5.1 Zufällige Körperbewegungen und Reflexe	62
1.5.2 Intentionale Körperbewegung und Ausdruck	63
1.6 Lip Synching als ‚so-tun-als-ob‘	66
1.7 Lip Synching als Täuschung	69
1.8 Lip Synching als Modulation einer Täuschung	71
1.9 Lip Synching im Musikclip als Darstellungshandlung	72
2 Musikclip	77
2.1 Die Entstehung des Musikclips und des Musikfernsehens	77
2.1.1 Vorgeschichte	77
2.1.2 MTVIVA	79
2.2 Definition(en) des Begriffs Musikclip und Abgrenzung zu verwandten Formaten	82
2.3 Kriterien zur Klassifikation von Musikclips	87
2.4 Ästhetik und Bauweise von Musikclips	92

3	Audiovision	99
3.1	Audiovision als natürliche Wahrnehmung	99
3.2	Systematische Ableitung der <i>Audio-Vision</i> im Musikclip	101
3.2.1	Von Face-to-Face zu filmisch tertiär-medial vermittelter Audiovision	101
3.2.2	Systematische Sonderfälle der Verbindung von Audio und Vision	104
3.2.3	Die spezielle <i>Audio-Vision</i> des Musikclips	113
3.3	Der gattungsspezifische Zusatzwert in der <i>Audio-Vision</i> des Musikclips	118
4	Materialverhältnisse	126
4.1	Live vs. recorded	127
4.2	Musikebene	129
4.3	Bewegtbildebene	137
5	Teilvisualisierung	146
5.1	Wechsel und Vermischung von Darstellungsformen	146
5.2	Musikbezogene Teilvisualisierung	152
5.3	Personenkonstellationen	165
6	Naturalisierung und Artifizialisierung	166
6.1	Naturalisierung und musikalisches Genre	168
6.2	Bildkategorien der Naturalisierung und der Artifizialisierung	174
6.3	Die Möglichkeit des Bildtons in der <i>Audio-Vision</i> des Musikclips	191
6.4	Formen der Manifestation des Bildtons	194
7	Strukturmodell	206
8	Lip Synching als Darstellungshandlung im Musikclip	213
8.1	Normalfall	214
8.1.1	Lip Synching und Bühnenperformance	216
8.1.2	Lip Synching und wechselnde Settings	224
8.1.3	Lip Synching und Illustration	229
8.1.4	Lip Synching und Plansequenz	233
8.1.5	Lip Synching und Gestik, Mimik und Physiognomie	236

8.2	Brüche	242
8.2.1	Alternation	243
8.2.2	Substitution	258
8.2.3	Objektivierung	287
8.2.4	„Falsches“ Lip Synching	310
8.2.5	„Unsichtbares“ Lip Synching	317
8.3	Gesang im Bildton und die Doppelung des Musiktons	319
8.4	Der Musikclip zu <i>A Song For The Lovers</i> als Sonderfall der <i>Audio-Vision</i>	332
Fazit		345
Literatur		354
Mediografie		367

Einleitung

Die vorliegende Arbeit widmet sich dem Phänomen des *Lip Synching*, *verstanden als Darstellung von Gesang innerhalb verschiedener Visualisierungszusammenhänge des populärkulturellen Gegenstands Musikclip*. Im Zentrum der Arbeit steht dabei die Frage nach der Etablierung und nach der Beschaffenheit der audiovisuellen Verhältnisse von Musikclips, die infolge des spezifischen Zusammenspiels von Bild, Text und Ton/Musik entstehen. Musikclips grenzen sich in ihrer Bauweise gegenüber verwandten audiovisuellen Medienprodukten (z.B. Werbung, Spielfilm) dadurch ab, dass einem bereits existierenden Musikstück nachträglich Bildereignisse hinzugefügt werden. Diese Tatsache ermöglicht insbesondere bezüglich der Darstellung der Musikaufführung ein zeitversetztes, nachträgliches Synchronisieren der Bildereignisse in Abstimmung mit den vorhandenen Musikereignissen. Die vorliegende Arbeit fokussiert somit weniger die präsentierten Inhalte oder mögliche bildbedingte Narrationen, sondern vielmehr die Synchronität zwischen zu sehendem Bildereignis und zu hörendem Musikereignis in Performancezusammenhängen sowie die daraus resultierende Vorstellung bzw. Wahrnehmung eines Musikclips als audiovisuelles Gesamtereignis.

Der Musikclip (bzw. das Musikvideo) war von Beginn an vor allem Gegenstand der film-, bild- und kunstwissenschaftlichen, der kultur- und sozialwissenschaftlichen und ebenso der (populär-)musikwissenschaftlichen Forschung. Über die Jahre hinweg lassen sich drei größere Forschungsfelder identifizieren, die jeweils fachspezifisch geprägt, aber ebenso kombinierbar sind: 1. die Analyse struktureller Zusammenhänge in und zwischen den drei Ebenen (Bewegt-)Bild, (Song-)Text und Musik(stück)/Ton im Sinne einer fokussierten Analyse spezifischer Bauweisen von Musikclips; 2. die (bild-)inhaltliche Analyse semantischer, symbolischer und generell ästhetischer Komponenten im Sinne exemplarischer Produktanalysen; 3. Rezeptionsanalysen im Sinne der (klassischen) Medienwirkungsforschung vor allem im Zusammenhang mit spezifisch jugendkulturellen Fragestellungen.

In der wissenschaftlichen Auseinandersetzung mit dem Medienprodukt Musikclip wurde zunächst insbesondere dessen eigenwillige Gestaltung der schnell geschnittenen Bildereignisse fokussiert, die sowohl als grelle Reizüberflutung wie auch als kunstvolle Inszenierung angesehen wurden. Die intermedialen Strukturen von Musikclips wurden in den frühen Produktanalysen noch kaum offen gelegt, stattdessen wurde dem damals neuartigen Medium hauptsächlich die Dominanz einer unstrukturierten Bilderflut unterstellt (vgl. Rauh 1985, S. 210).

Die spezifische Semantik und die rhythmischen Zusammenhänge zwischen Musikstück und Bewegtbild wurden erst mit der steigenden Verbreitung und Anzahl an Musikclips erkannt und fokussiert (vgl. Jacke 2003, S. 29ff.). Anfängliche Studien suchen sich hierfür als Analysegegenstand des Öfteren noch bekannte Extremformen des Musikclips (z.B. jenen zum Song *Thriller* von MICHAEL JACKSON), die dann als Paradebeispiel der jeweils vertretenen Forschungshaltung fungieren (vgl. Altrogge 2002, S. 2441ff.). Eine wesentliche Erkenntnis der eher (kultur-)kritisch angelegten Studien der Anfangszeit ist jedoch, dass der musikindustrielle Kontext samt den zugehörigen Begleittexten immer in die Analyse von Musikclips mit einzubeziehen ist, da nur so das Gesamtphänomen hinsichtlich seiner imagebildenden Funktion fassbar gemacht werden kann (vgl. Pape/Thomsen 1997, S. 210f.). Aufgrund der häufigen Einzelfallstruktur der Analysen und der spezifischen Fragestellungen an das Produkt Musikclip bleiben die strukturellen Beziehungen von Musik und Bild zu Beginn der wissenschaftlichen Auseinandersetzung allerdings weitestgehend im Dunkeln (vgl. Altrogge 1994, S. 196f.). Im Laufe der Zeit gewinnt die Analyse von Musikclips unter Einbezug der spezifischen medialen Rahmenbedingungen an interdisziplinärer Schärfe, wenn auch zunächst nur wenige thematische Forschungsstränge entstehen. In Inhaltsanalysen standen von Beginn an die potenziellen Problemkontexte verschiedener Arten gewaltaffiner und sexualisierter Darstellungen im Vordergrund, sowie Fragen bezüglich der Geschlechterkonstruktion, der Genderverhältnisse und ethnischer und rassistischer Darstellungsweisen in Musikclips verschiedener Typen und Genres (vgl. Neumann-Braun/Mikos 2006, S. 25ff.). Eine wesentliche Erkenntnis dieser frühen Musikclipanalysen ist allerdings, dass Musikclips als audiovisuelle Medienprodukte über eigenständige Erzählformen mit oftmals hohen künstlerischen Ansprüchen verfügen, die sich aus der spezifischen Kombination von vorhandenem Musikstück und Songtext mit hinzugefügten Bilderereignissen ergeben. Entsprechend komplex konzipierte und ästhetisch herausstechende Werke werden in der Folge oftmals (bild-)inhaltlich wie auch kulturell positiv rezipiert und bewertet. Gerade mit Blick auf Produktanalysen muss jedoch beachtet werden, dass sich Musikclips immer einer musikalischen Grundlage bedienen. Wird dies vergessen oder ausgeblendet, können insbesondere Analysen so genannter Konzept- oder Art-Clips zu einer Überbewertung des Visuellen und einer Vernachlässigung der zugrunde liegenden Komponenten Musikstück und Songtext führen (vgl. Neumann-Braun/Schmidt 1999, S. 16f). Allgemein kann festgehalten werden, dass es sich bei dem Format Musikclip und den dortigen Zusammenhängen von Musikstil und Bildinhalten stets um ein wandelbares Phänomen handelt, welches es demzufolge in Analysen immer wieder zu aktualisieren gilt: Zum einen, weil neu entstehende (und auch wieder belebte) Musikstile in Abgrenzung zu bestehenden Musikstilen eine ebenso neue

Symbolik und Ästhetik entwickeln können, die gleichermaßen durch intermediale Bezüge geprägt werden; zum anderen, weil innerhalb der Zitativität des Musikclips ironisierende Zusammenhänge, Adaptionen und Formen von Pastiche bedacht werden müssen. So lässt sich erkennen, dass bestimmte musikalische Genres eigene, (jugend-)kulturell relevante Visualisierungsformen entwickelt haben, die im Musikclip in Verbindung mit populärer Musik reproduziert werden können: z.B. kehren klassische Heavy-Metal-Clips und auch Rap/Hip-Hop-Clips immer wieder zum Konzept der pseudo-dokumentarischen Musikperformance zurück (vgl. Altrogge 2001b, S. 20f.), Techno- und Dance-Clips arbeiten bereits früh mit Computeranimationen (vgl. Pape/Thomsen 1997, S. 210).

Die vorliegende Arbeit widmet sich allgemein den gattungsspezifischen Konstruktionsprinzipien und der Struktur der *Audio-Vision* von Musikclips und fragt im Sinne der *Rahmen-Analyse* (1977) nach Erving Goffman „Was geht hier eigentlich vor?“ (S. 16). Das heißt, welche Bezugnahmen sind zwischen Bild und Ton nötig, um die Bild-Text-Ton-Verhältnisse des Musikclips als zusammengehörige *audio-visuelle* Ereignisse präsentieren und verstehen zu können. Kernaspekt ist hierbei, dass der Musikclip als tertiär-medial vermitteltes *audio-visuelles* Produkt vorliegt und als solches dadurch gekennzeichnet ist, dass Bild- und Tonereignisse voneinander *entkoppelt* und anschließend wieder miteinander *verkoppelt* werden können. Diese Tatsache kommt den musikbezogenen Darstellungen in Musikclips insofern zu Gute, als somit z.B. ermöglicht wird, dass die nachträgliche musikbezogene Performance einer beliebigen Person oder Figur mit dem vorab erzeugten zu hörenden musikalisch-klanglichen Element der Interpretation verkoppelt werden kann.

Als Analysegegenstand *audio-visueller* Bezugnahmen eignen sich folglich insbesondere Performance-Clips, d.h. Musikclips, in denen zur Gänze oder in Teilen die Aufführung des zu hörenden Musikstücks oder einzelner Elemente des Musikstücks von Personen oder Figuren dargestellt wird. Die Darstellung der Musikperformance konstruiert ein besonders starkes (Abhängigkeits-)Verhältnis in der Synchronisation von existierendem Musikstück und hinzugefügten Bildereignissen, da sich letztere direkt auf die vermeintliche (Re-)Produktion der Musik und deren Visualisierung beziehen. Ohne das zugrunde liegende Musikstück verlieren die Performancedarstellungen eines Musikclips ihre Sinnhaftigkeit, da ihnen die akustische Referenzgröße fehlt. Für die synchrone Kombination mit dem Musikstück orientieren sich Performancedarstellungen hingegen direkt am zugrunde liegenden musikalisch-klanglichen Material sowie an bekannten Live-Darbietungen eines Musikstücks. Auf diese Weise konstruieren und vermitteln Performance-Darstellungen in Musikclips Vorstellungen zusammengehöriger, audiovisueller Gesamt ereignisse. Die Nachahmung des Gesangs durch die Darstellung des Lip Synching nimmt in Performance-Clips eine

besondere Stellung ein, denn das Lip Synching visualisiert mehrere Aspekte zugleich: erstens das musikalisch-klangliche Element Gesang eines Musikstücks; zweitens die Stimme als individuelles, körpereigenes musikalisches Instrument; drittens den Akt des Singens als das ‚Spielen‘ des Instruments Stimme; und viertens den (menschlichen) Körper als anatomischen Klangkörper, durch den Stimme und Gesang aktiv produziert werden. Die spezifische Grundlage des Lip Synching im Musikclip ist dabei, dass in der ursprünglichen, vorangegangenen Gesangsproduktion, die jetzt als Teil des Musikstücks zu hören ist, das ‚gespielte‘ Instrument Stimme und die Person (der/die InterpretIn), die das Instrument Stimme ‚spielt‘, auf ein und denselben Körper zusammenfallen (den der InterpretIn).

Wendet man sich speziell dem Lip Synching in Musikclips zu, so ist festzustellen, dass dieses Phänomen in der bisherigen Musikclipforschung bis dato nicht detaillierter thematisiert worden ist. Ein Grund hierfür liegt möglicherweise darin, dass der Begriff ‚Lip Synching‘ z.B. eher als die Vortäuschung von Gesang in Playback-Auftritten oder als ein gängiges Element in Musikfilmen verstanden wird/wurde (vgl. Goodwin 1992, S. 30), jedoch kaum als zentrale musikbezogene Darstellungsform im audiovisuellen Kurzformat Musikclip. So führt beispielsweise Andrew Goodwin (1992) in *Dancing in the distraction factory* das Lip Synching als ein Darstellungselement an, das im Kontext von Fernsehauftritten von dem/der SängerIn einer Rock-/Pop-Band scheinbar beherrscht werden muss (vgl. Goodwin 1992, S. 9), das aber ebenso aufgrund der exakten Nachstellung des vermeintlichen Singens stets Kontroversen und Diskussionen über die ‚Echtheit‘ einer musikalischen Performance auslöst (vgl. Goodwin 1992, 32ff.). Das Lip Synching fungiert in diesem Zusammenhang als ein Unterscheidungsmerkmal zwischen den basalen Musikrichtungen Rock und Pop, denn oftmals ist die Nachahmung des Gesangs lediglich ein Element innerhalb einer Pop-Performance, bei der das gesamte Musikstück von einem Tonträger abgespielt und folglich die gesamte musikalisch-klangliche Live-Produktion vorgetäuscht wird (vgl. Goodwin 1992, S. 36). Hinsichtlich des Musikclips beschränkt sich Goodwin nur auf beiläufige Erwähnungen des Lip Synching in der Deskription der Bildereignisse einzelner Performance-Clips, die das Lip Synching als besonders hervorstechendes visuelles Gestaltungsmittel zur Inszenierung der Star-Person nutzen (vgl. Goodwin 1992, S. 64 und S. 94). Eine dezidierte Auseinandersetzung mit dieser nahezu allgegenwärtigen musikbezogenen Darstellung in Musikclips bleibt in Goodwins frühem Werk der Musikclipforschung somit aus. Dies ist insofern bemerkenswert, als der Musikclip (bzw. das Musikvideo) von Beginn an darauf ausgerichtet war, den/die (Musik-)KünstlerIn in Ergänzung zu Live-Performances sowie den (Pop-)Song als bereits existierendes Klangmaterial über das reine Hörerlebnis hinaus in einem zusätzlichen und eigenständigen audiovisuellen Zusammenhang zu präsentieren. Dem Lip Synching

kommt dabei innerhalb dieser Form der Performance eines (Pop-)Songs eine zentrale Rolle zu, denn es visualisiert den Akt des Vortragens des Songtexts. Das Lip Synching zeigt darüber hinaus an, welcher der zu sehenden (Klang-)Körper ‚singt‘ bzw. ‚erklingt‘ und zudem, welche Person den zu hörenden Songtext im Sinne eines Erzählers in den Bildereignissen vorträgt bzw. präsentiert. Railton/Watson (2011) weisen in *Music Video and the Politics of Representation* auf die grundsätzliche Artifizialität der Performancedarstellungen in Musikclips hin und führen das Lip Synching als einen der visuellen Bausteine sogenannter „staged performance music videos“ (S. 58) an, d.h. als im Grunde notwendiges Element einer einzig für den Musikclip bzw. für die Musikclipproduktion inszenierten Musikperformance. Ebenso betonen Railton/Watson den Aspekt der physischen Darstellung der rhythmischen Elemente des Musikstücks durch Tanzperformances, aber auch durch die Nachstellung des Gesangs mittels Lip Synching (vgl. Railton/Watson 2011, S. 58ff.). Das Lip Synching wird folglich als ein Kernelement in der ‚staged performance‘ angesehen (vgl. Railton/Watson 2011, S. 62), eine tiefer gehende Analyse hinsichtlich der *audio-visuellen* Zusammenhänge, die durch diese Form der nachgestellten Musikdarbietung entstehen, bleibt jedoch auch bei Railton/Watson aus. Insgesamt ist über diese Beispiele hinaus zu erkennen, dass dem Phänomen Lip Synching in der bisherigen wissenschaftlichen Auseinandersetzung mit Musikclips keineswegs jene Wichtigkeit beigemessen wurde, die es als *audio-visuelles* Konstrukt in sich birgt. Hinsichtlich der Klassifikation von Musikclips betont beispielsweise Altrogge (2001b) in seiner interdisziplinären Studie *Tönende Bilder* zwar den Appellcharakter des gesungenen Texts und die auf den/die SängerIn fokussierte Inszenierung (vgl. Altrogge 2001b, S. 26ff.), jedoch findet auch bei Altrogge das Lip Synching als entsprechend zentrales Darstellungselement einer nachgestellten bzw. inszenierten Performance eines Songs keine Erwähnung.

Unabhängig von der Frage nach der Klassifikation von Musikclips ist jedoch zu erkennen, dass die singuläre Nachahmung des Gesangs bzw. des ‚Singens‘ durch eine Person oder Figur der Bildereignisse eine häufige – wenn nicht die häufigste – musikbezogene Darstellung in Musikclips ist. Alleine diese elementare Tatsache ist Grund genug für eine dezidierte Auseinandersetzung mit dieser Darstellungsstrategie. Im Zusammenhang mit dem Aspekt der Personen und Figuren wird das Lip Synching auch in zahlreichen Analysen erwähnt, vor allem hinsichtlich der Zuordnung zusätzlicher Erzählebenen der Bildereignisse zu den Schilderungen des Songtexts (vgl. z.B. Railton/Watson 2011, S. 8f.), und wenn darüber hinaus basale Fragen nach einem Gesamtnarrativ aufgeworfen werden, welches sich aus dem Lip Synching einer Person/Figur und weiteren, nicht performativ umgesetzten Visualisierungen des Songtexts ergibt (vgl. z.B. Railton/Watson 2011, S. 57f.). So diskutiert beispielsweise Carol Vernallis (2004) in *Experiencing Music Video* den Aspekt, dass eine Person im Musikclip durch ihr

Lip Synching in verschiedenen Zusammenhängen als ‚erzählende Instanz‘ ausgewiesen wird. Vernallis zufolge kann das Lip Synching einer (Star-)Person als Element in narrativ konzipierten Musikclips z.B. unterschiedlich starke bzw. dichte Verbindungen zwischen den Songtextschilderungen und einzelnen Situationen der Bildereignisse herstellen. Das Lip Synching erzeugt mitunter den Eindruck, dass der/die SängerIn den Songtext im Kontext des Bildgeschehens ‚erzählt‘/‚erzählen‘ würde. Vernallis weist allerdings zurecht darauf hin, dass derartige ‚Erzählsituationen‘ sowohl durch die Mittel der Darstellung (vor allem durch Montage) sowie durch Änderungen innerhalb von Raum und/oder Zeit der dargestellten Ereignisse relativ schnell durchbrochen und dekonstruiert werden können (vgl. Vernallis 2004, S. 9). Anhand der Analyseebene ‚Person/Figur‘ kommt Vernallis folglich zu der Einschätzung, dass eine visuelle Anwesenheit der InterpretIn(nen) in den Bildereignissen die Darstellung des Lip Synching nahezu unumgänglich einfordert und somit im Verhältnis von zu hörender und zu sehender Person eine Darstellungskonvention entstanden ist. So formuliert Vernallis bereits in der Einleitung zu *Experiencing music video*:

„The presence of star performers points to the strictest and most pervasive of music video’s conventions: a video must provide a flattering depiction of the singer lip-syncing the song“ (Vernallis 2004, S. XII).

In ihren Analysen zeigt Vernallis dementsprechend auf, dass die gesamt-narrativen Strukturen von Musikclips über die Personen-/Figurenverhältnisse und mitunter über die Analyse der Darstellung des Gesangs aufgeschlüsselt werden können, wodurch sich die spezifische Konstruktion der *Audio-Vision* von Musikclips bereits andeutet. Besonders hervorzuheben sind dabei laut Vernallis Musikclips, die z.B. Hauptcharaktere, welche nicht mit den MusikerInnen übereinstimmen, ausschließlich in Situationen zeigen, in denen sie nicht sprechen/singen können (vgl. Vernallis 2004, S. 5), oder Musikclips, in denen Personen/Figuren (z.B. Statisten) in den Bildereignissen als prinzipiell ‚stumm‘ ausgewiesen werden (vgl. Vernallis 2004, S. 30). Die spezifischen audiovisuellen Verkoppelungen, die sich im Musikclip durch das Zusammenspiel aus Bewegtbild, (Song-)Text und Musik/Ton – und somit auch durch das Lip Synching – ergeben, verdeutlicht Vernallis anhand der filmwissenschaftlichen Grundlage des ‚sync point‘ auf Basis von Michel Chions (1994) *Audio-vision. Sound on screen*. Vernallis betrachtet dabei vor allem die Möglichkeiten der Verstärkung musikalischer Parameter und Strukturen durch entsprechende synchrone (Klang-)Bildereignisse, jedoch findet das Phänomen des Lip Synching als entsprechend synchrones und bewusst herbeigeführtes Zusammentreffen von Gesang als musikalischem Ereignis und Nachstellung des Singens als visuellem Ereignis keine Erwähnung. Hinsichtlich der ‚sync points‘ sind in der Folge vor allem Abweichungen in der Verkoppelung von Lip Synching und darstellender Person/Figur von Interesse. Beispielsweise greifen Railton/Watson den Aspekt des Bruchs zwischen zu sehender und

zu hörender Person auf, der durch das Lip Synching einer ‚falschen‘ Person in den Bildereignissen entsteht, und setzen diesen in Bezug zur Herstellung einer genderkorrekten, authentischen Präsentation der (Star-)Person(en) innerhalb des Gesamtzusammenhangs des Musikclips (vgl. Railton/Watson 2011, S. 76f.). Auch Vernallis weist in ihrer kurzen Analyse des Musikclips zum Song *Gett Off* von PRINCE & THE NEW POWER GENERATION auf das Lip Synching durch Personen, die nicht die InterpretInnen des Musikstücks sind, hin: „(...) here performers lip-sync and play instruments different from those that occurred on the recording“ (Vernallis 2004, S. 180).

Wendet man sich folglich der Analyse von Musikclips und vor allem der Frage nach der Konstruktion der *Audio-Vision* zu, so sind diese einführend erläuterten Erkenntnisse über die Bauweise von Musikclips und ihre spezifischen Bild-Musik-Relationen entsprechend grundlegend. Die Komplexität des *audio-visuellen* Zusammenspiels in Musikclips ist jedoch bisweilen schwer zu erfassen und darzustellen, denn Musikclipsanalysen jeder Art haben nach wie vor das zentrale Problem einer fehlenden adäquaten Methode, welche die drei Teilbereiche (Bewegt-)Bild, (Song-)Text und Musik(stück)/Ton gleichermaßen und im Sinne einer interdisziplinären Ausrichtung gleichberechtigt integriert. Für die Bildebene werden oftmals Methoden der Film- und/oder Fernsehanalyse oder der Bildanalyse (Fotografie) und Ikonografie im Allgemeinen herangezogen, wodurch in der Folge musikalisch-klangliche und (song-)textliche Analysekategorien fast erwartungsgemäß vernachlässigt werden. Klassische musikwissenschaftliche Ansätze sind hingegen nur sehr selten in der Lage, die populärmusikalischen Strukturen des Musikclips gegenstandsangemessen greifbar zu machen, geschweige denn das Bildgeschehen in seiner Komplexität zu erfassen (vgl. Goodwin 1992, S. 20). Dies führt letztlich dazu, dass inzwischen eine Vielzahl von Vorschlägen und Ansätzen zur wissenschaftlichen Analyse von Musikclips existiert, so dass die eigentlich zu vermeidende Kontingenz durch die Komplexität und zum Teil durch die Unvereinbarkeit der vorgeschlagenen Analysekategorien erneut erzeugt wird (vgl. Helms 2003, S. 99). Insbesondere die adäquate Integration der musikalisch-klanglichen Elemente des zugrunde liegenden Musikstücks bleibt oftmals außen vor, wodurch nicht selten eine detaillierte Analyse des Bildgeschehens der reinen Deskription der Musikereignisse gegenübergestellt wird. Grundlegend ist festzuhalten, dass die Analyse von Musikclips immer unter der Prämisse der Integration methodischer Ansätze zur Analyse aller Ebenen ((Bewegt-)Bild, (Song-)Text und Musik(stück)/Ton) erfolgen sollte, so dass der Musikclip auch bei darauf aufbauenden spezifischen Analysen einzelner Ebenen in seiner Grundstruktur zugänglich gemacht wird (vgl. Jost/Klug 2009, S. 197f.). Diese Notwendigkeit lässt sich bereits in den zuvor erläuterten, jedoch spärlichen Erwähnungen des Phänomens Lip Synching erkennen, denn die Tatsache, dass z.B. die Möglichkeit besteht, eine Person beim ‚Singen‘ zu zeigen,

die nicht mit dem/der zu hörenden SängerIn übereinstimmt, bedingt sowohl die singuläre Auseinandersetzung mit dem Bild- und dem Musikereignis, als auch die Analyse der *audio-visuellen* Verkoppelung beider zu einem präsentierten Gesamtereignis.

Gerade im Zusammenhang mit der Analyse der angesprochenen ‚sync points‘, ist insbesondere der Aspekt der Visualisierung von Bild-Text-Ton/Musik-Bezügen in entsprechenden Analyseansätzen¹ zu betonen. Mit Hilfe schematischer Darstellungen in Form einer Partitur oder eines Transkripts können im *audio-visuellen* Zusammenspiel z.B. nachträgliche Performancedarstellungen im Verhältnis zu den zugrunde liegenden musikalisch-klanglichen Elementen verortet und deren Struktur aufgezeigt werden. Ebenso lässt sich hinsichtlich des Lip Synching die Verkoppelung von Bildereignis und musikalisch-klanglichem Ereignis visualisieren, so lassen sich unter anderem Konstruktionen wie die erläuterte Verkoppelung der Gesangsdarstellung einer dritten Person mit der erklingenden Stimme der wirklichen SängerIn markieren. Insbesondere bei der Analyse von Musikclips, die unterschiedliche Visualisierungsstrategien miteinander verknüpfen (z.B. das Lip Synching als Teil-Performance innerhalb narrativer Bildereignisse), ist der Einbezug entsprechender computergestützter Analyseprogramme² von Vorteil, da mit deren Hilfe die mitunter ständig wechselnden

- 1 Beispielsweise erstellt Rappe (2010) eine mehr als 50 Seiten lange „audiovisuelle Partitur“, in der die musikimmanenten Bedeutungsstrukturen des Musikclips zu *Work It* der US-amerikanischen Rapperin MISSY ELLIOT in einer detaillierten Bild-Text-Ton-Analyse transkribiert sind (vgl. Rappe 2010b). Rappe untersucht aus kulturwissenschaftlicher Perspektive das „signifying“ und die Verschränkungen der lyrischen, visuellen und musikalischen Elemente des afroamerikanischen Hip-Hop (vgl. Rappe 2010a). Die ambitionierte und umfangreiche Analyse des durch visuelle und songtextliche Samples verschachtelten Songs *Work It* zeigt sowohl die Notwendigkeit als auch die Probleme einer derart feingradigen Bild-Text-Ton-Analyse auf. Eine vollständige Synchronität zwischen Notation und einzelнем Bildereignis sowie zwischen teilweise zweizeiligen Transkriptionen des Songtexts und dem entsprechenden Bildereignis kann in dieser Printdarstellung nicht erreicht werden. Die Detailliertheit der Transkription derartiger Bezugnahmen vermag es jedoch, die Konstruktion der *Audio-Vision* des Musikclips anhand entsprechender musikbezogener Performancedarstellungen aufzuzeigen.
- 2 Eine vollkommen synchrone Transkription der mannigfachen Bild-Text-Ton/Musik-Bezüge von z.B. Musikclips erscheint letztlich nur in computergestützten Analysen möglich. Dies gilt ebenso für die Frage der adäquaten Visualisierung audiovisueller Zusammenspiele. Mit diesem Anspruch wurde am Seminar für Medienwissenschaft der Universität Basel die Web-Applikation „trAVis“ (www.travis-analysis.org) entwickelt. „trAVis“ ist ein musikzentriertes Transkriptionsprogramm für audiovisuelle Medienprodukte, welches eine gegenstandsangemessene, computergestützte Transkription und Analyse musikbasierter, audiovisueller Medienprodukte ermöglicht. Die Web-Applikation verbindet bildbezogene und texthermeneutische Ansätze mit musikwissenschaftlichen Zugängen. Eine frei editierbare Arbeitsfläche ermöglicht die gleichzeitige schematische Darstellung aller transkribierten Bild-, Text- und Tonereignisse (vgl. Jost/Klug/Schmidt/Reautschnig/Neumann-Braun 2013).

Bezugnahmen zwischen hinzugefügten Bildereignissen und zugrunde liegendem Musikstück dezidiert aufgezeigt und folglich die Konstruktionsprinzipien der *Audio-Vision* von Musikclips offengelegt werden können.

Auf Grundlage der erläuterten Bauweise von Musikclips und der genannten Notwendigkeit einer gegenstandsangemessenen Methode, zeigt die vorliegende Arbeit auf, wie die Konstruktionsprinzipien der *Audio-Vision* von Musikclips mittels des Phänomens Lip Synching zugänglich gemacht und analysiert werden können. Wie zuvor erläutert, kann die Darstellung des Lip Synching als universelles *audio-visuelles* Phänomen in Musikclips verstanden werden, da es sich um die häufigste musikbezogene Darstellung handelt, die in der tertiär-medialen Vermittlung zugleich eine einzigartige Verbindung von Klang und Körper konstruiert. Das Phänomen Lip Synching bietet somit einen äußerst lohnenswerten Zugang zur Analyse der mannigfachen Bezugnahmen zwischen Bildereignissen und Musikstück, welche die spezifische *Audio-Vision* des Musikclips konstruieren. Anhand der Verkoppelung der nachträglichen Gesangsdarstellung mit der klanglichen Beschaffenheit des vorliegenden Gesangsparts lassen sich in der Folge die manifesten wie auch die latenten Bezugnahmen zwischen Bildereignissen und Musikereignissen bestimmen.

Hieraus ergibt sich ein Aufbau entlang der folgenden Kernaspekte:

- Grundlegend: die physische Verbindung von sichtbarem Körper und hörbarer Stimme und die Wahrnehmung von medial unvermittelter Audiovision, die demnach als ‚natürliche‘ Audiovision verstanden werden kann;
- daraus abgeleitet: die Möglichkeiten der *Entkoppelung* und der (erneuten) *Verkoppelung* von Bewegtbild (i.d.F. den Lippenbewegungen als Element des Gesamtgeschehens der Bildebene) und Ton/Musik (i.d.F. des Gesangs als Element des Musikstücks) innerhalb tertiär-medial vermittelter Audiovision;
- speziell: die Varianten der Darstellung des Gesangsparts bzw. die Nachahmung des ‚Singens‘ mittels Lip Synching im Kurzformat Musikclip sowie die Rolle des Lip Synching in der Konstruktion der gattungsspezifischen *Audio-Vision* des Musikclips.

Im Zentrum steht dabei die Entwicklung eines *Strukturmodells des Bezugssystems in der Audio-Vision des Musikclips* (Kap. 7), welches die bereits angedeuteten Bezugnahmen zwischen den Bild- und Ton-/Musikkomponenten systematisiert. Dieses Strukturmodell findet seine Anwendung in den anschließenden Musikcliplanalysen (Kap. 8), die das Bezugssystem anhand der musikbezogenen Darstellung des Lip Synching als besonderer Form des Zusammenwirkens von Bewegtbildereignis (Lippenbewegung) und musikalisch-klanglichem Element (Gesangspart) ausführlich veranschaulichen.

In Kapitel 1 der vorliegenden Arbeit wird Lip Synching zunächst grundlegend als audiovisuelles Phänomen hergeleitet und darauf folgend im Speziellen das Lip Synching in Musikclips gegenüber verwandten Formen der tertiär-medialen (Neu-)Kombination von Bild- und Ton-/Musikereignissen abgegrenzt. Als Basis für die späteren Musikclipanalysen wird in Kapitel 1 eine formatspezifische Definition des Lip Synching in Musikclips formuliert. Dabei wird deutlich, dass das Lip Synching im Musikclip als intentionale und somit bewusst ausgeführte Handlung zu verstehen ist. Dies wird unter Bezug auf die *Rahmen-Analyse* nach Erving Goffman (1977) dargelegt, so dass sich das Lip Synching anhand der Rahmen der Täuschung und des ‚so-tun-als-ob‘ verorten lässt.

Anschließend wird in Kapitel 2 die der zentrale Analysegegenstand Musikclip historisch verortet sowie in seiner spezifischen Bauweise und Ästhetik dargelegt und als eigenständige audiovisuelle Gattung definiert, zudem werden die Möglichkeiten der Klassifikation dieses Kurzformats diskutiert.

Auf Basis der Verortung des Phänomens Lip Synching liegt in Kapitel 3 der Fokus auf der Diskussion des Begriffs der Audiovision sowie auf der sukzessiven Herleitung der spezifischen Form der *Audio-Vision* des Musikclips. Hierbei ist von Audiovision im Sinne ‚natürlicher‘, d.h. nicht medial vermittelter Wahrnehmungszusammenhänge auszugehen und auf Formen filmischer, d.h. tertiär-medial vermittelter Audiovision zu schließen. Als Ergebnis von Kapitel 3 werden die theoretischen Konzepte des Lip Synching und der *Audio-Vision* als Grundlage für die anschließenden Beispielanalysen zusammengeführt.

Kapitel 4 widmet sich in einem nächsten Schritt den Materialverhältnissen des Analysegegenstands Musikclip, welche die spezifische Form der *Audio-Vision* letztlich erst ermöglichen. Das Zusammenwirken von Bewegtbild und Musikstück ist im Musikclip dadurch gekennzeichnet, dass ein bereits eigenständig existierendes Musikstück nachträglich durch verschiedenartige Bilderfolgen ergänzt wird, die in ebenso unterschiedlicher Weise auf die musikalisch-klangliche Struktur des Musikstücks und/oder auf die Inhalte des Songtexts Bezug nehmen (können). Dies ist für den Aspekt der Entkoppelung von Bild und Musik und der anschließenden Verkoppelung zu einem *audio-visuellen* Gesamt ereignis entsprechend zentral.

Aufbauend auf den Materialverhältnissen des Musikclips und den Bezügen zwischen Bildereignis und Musikereignis verdeutlicht Kapitel 5, dass es sich bei der Darstellung des Lip Synching lediglich um die Visualisierung des Gesangsparts als eine Form der Teilvisualisierung des zugrunde liegenden und erklingenden Musikstücks handelt. Hierin liegt ein entscheidender Faktor für die Möglichkeiten der (temporären) Verkoppelung von Bildereignis (Lippenbewegung) und Musikereignis (Gesang bzw. Stimme), die im Weiteren in der Analyse der Beispielclips (Kap. 8) aufgezeigt werden.

Kapitel 6 behandelt in der Folge den Kernaspekt der Annäherung an ein imaginierbares, ‚natürliches‘ audiovisuelles Ereignis (d.h. an Formen der nicht medial vermittelten Präsentation eines Musikstücks durch den/die SängerIn bzw. die Band) in der medialen (Wieder-)Verkoppelung von Bild und Ton im Musikclip. Für die Darstellung des Lip Synching ist dabei die Komponente der Person/Figur ausschlaggebend, d.h. die basale Frage der Übereinstimmung (Naturalisierung) oder des Bruchs (Artifizialisierung) zwischen der zu sehenden, das Lip Synching ausführenden Person, und der gleichzeitig zu hörenden Person, die den Gesangspart des erklingenden Musikstücks vorgängig produziert hat. In Kapitel 6 werden die Kriterien der Naturalisierung und der Artifizialisierung zur Analyse möglicher Brüche in der Konstellation *audio-visueller* Zusammenhänge für die Bildebene wie auch für die Tonebene des Musikclip hergeleitet. Hierbei sind die Möglichkeiten der Manifestation von zusätzlichem Bildton (Tonereignisse, die dem Bildgeschehen und nicht dem Musikstück zuzuordnen sind) zentral, da dieser zusammen mit dem zugrunde liegenden Musikstück oder anstelle von diesem zu hören sein kann. Für den Fall derartiger akustischer Ergänzungen kommt dem Lip Synching in Musikclips erneut eine entscheidende Rolle zu, da es als Darstellungshandlung lediglich den sichtbaren Part der *Audio-Vision* beisteuert. Der zu hörende Gesang ist hingegen Teil des bereits materialisiert vorliegenden Produkts Musikstück bzw. Tonträger. Jedoch obliegt dem Lip Synching als Darstellung des ‚Singens‘ in der Bildebene des Musikclips ebenso die Möglichkeit, zu eigentlichem Gesang zu werden, d.h. einen entsprechenden Gesangston zu erzeugen, der in der *Audio-Vision* des Musikclips zu manifestem, musikstückfremdem Bildton werden kann. Das Lip Synching kann folglich in bestimmten Fällen im Musikclip als audiovisuelles Ereignis vorliegen.

Als Ergebnis der Diskussion der Materialverhältnisse und der dadurch bedingten spezifischen Form der *Audio-Vision*, wird in Kapitel 7 schließlich das *Strukturmodell des Bezugssystems in der Audio-Vision des Musikclips* formuliert und entsprechend erläutert. Dieses Strukturmodell verdeutlicht in universeller Form die Bezüge zwischen manifesten und latenten Komponenten der Bild- und der Musikebene des Musikclips.

Das entwickelte Strukturmodell dient folglich in Kapitel 8 als Grundlage für die Analyse der Formen des Lip Synching als Darstellungspraxis in unterschiedlichen Visualisierungskonzepten ausgewählter Musikclips. Getreu der erläuterten Unterscheidung in naturalisierende und artifizialisierende Komponenten, werden zunächst Normalfälle der Darstellung des Gesangs angeführt. Ausgangspunkt ist die ‚reine‘ Bühnenperformance des Songs, von der aus schrittweise Abstraktionen identifiziert werden können. Demgegenüber werden die artifizialisierenden Darstellungsstrategien als Brüche hinsichtlich der Vergleichsfolie einer als ‚natürlich‘ empfundenen Audiovision verstanden. Der Aspekt des Zusammenstreffens von visueller Person/Figur (in der Bildebene) und auditiver Person bzw.

InterpretIn (in der Musikebene bzw. im Musikstück) ist dabei für die Beispielclips beider Grundausprägungen zentral. Durch die sukzessive Artifizialisierung der Person/Figur und letztlich des (Klang-)Körpers wird folglich der Status des Lip Synching als Darstellungshandlung in Musikclips erörtert und die spezifische Form der *Audio-Vision* in Musikclips dargelegt.